BigData Express

– Toward Schedulable, Predictable, High-performance Data Transfer

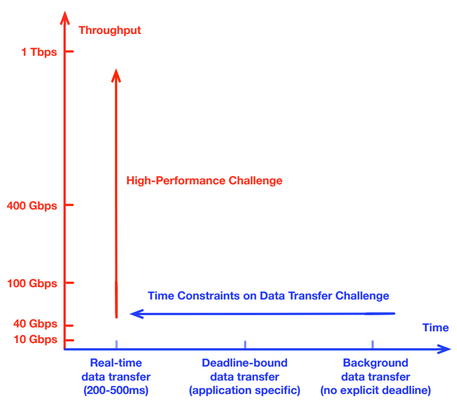
Q. Lu*, S. Sasidharan*, L. Zhang*, W. Wu*, P. DeMar*, J. Wang+, G. Liu+, N. Rao+, S. A. R. Asif#, J. Kim#, S. Noh#

* Fermilab, + Oak Ridge National Laboratory, #KISTI

Motivation

DOE Data Transfer Challenges

- High-performance challenges
- Time-constraint challenges
  - Real-time data transfer
  - Deadline-bound data transfer
  - Background data transfer



Data Transfer – State of the Art

- Advanced data transfer tools and services developed
  - GridFTP, BBCP, PhEDEx, LIGO Data Replicator, Globus Online
- Numerous enhancements
  - Parallelism, Science DMZ architecture, Terabit networks

Existing data transfer tools and services will NOT be able to successfully address the challenges of data transfer to support extreme-scale science applications

Problems with Existing Data Transfer Tools and Services

- Disjoint end-to-end data transfer loop
- Cross-interference between data transfers
- Oblivious to user requirements (e.g., deadlines and Qos)
- Inefficiencies arise with existing data transfer tools on DTNs

Our Solution

The BigData Express Project

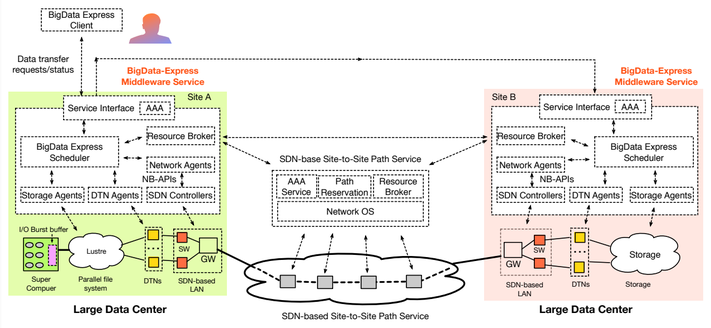
- Collaborative effort by Fermilab and Oakridge National Laboratory
- Funded by DOE's Office of Advanced Scientific Computing Research (ASCR)
- A three-year research project
- http://bigdataexpress.fnal.gov

BigData Express seeks to provide a schedulable, predictable, and high-performance data transfer service for DOE's large-scale science computing facilities (e.g., LCFs, US-LHC computing facilities)

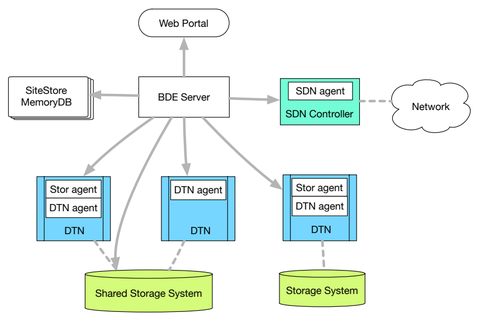
BigData Express Key Features

- A data-transfer-centric architecture to seamlessly integrate and efficiently coordinate the various resource in an end-to-end loop
  - Directly schedule various local resources within a site
  - A distributed rate-based resources brokering mechanism to coordinate resources across sites
- A time-constraint-based scheduler to schedule data transfer tasks
- An admission control mechanism to provide guaranteed resources for admitted data transfer tasks
- An end-host-based rate control mechanism to improve data transfer schedulability and reduce cross-interference between data transfers
- Extensive use of SDN and SDS to improve network and storage I/O performance

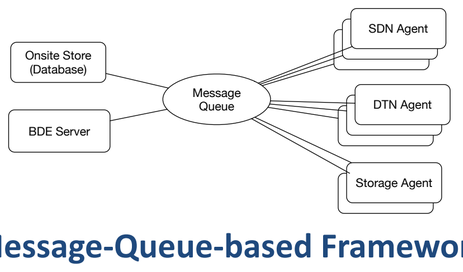
BigData Express Design & Architecture



BigData Express Architecture (I)

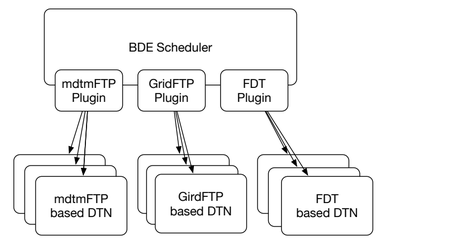


BigData Express Architecture (II)

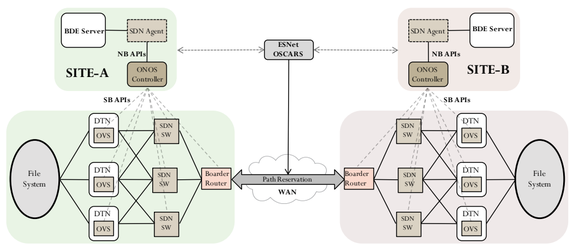


Message-Queue-based Framework

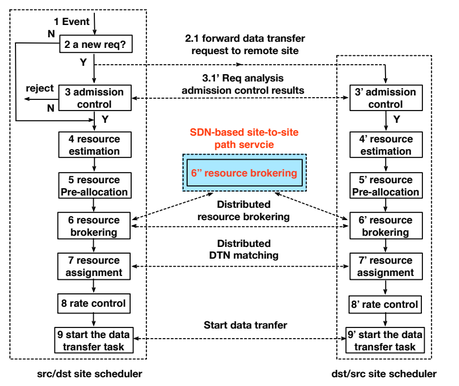
- Unified APIs
- Extensible



Pluggable Open Architecture



End-to-End Network Path



How does BigData Express Work?

How does BigData Express Work?

- The BigData Express scheduler implements a time-constraint-based scheduler to schedule resource for data transfer tasks
- Each resource will be estimated, calculated, and converted into a rate that can be apportioned to data transfer tasks
- On an event-driven or periodic basis, the scheduler will perform the following tasks:
  - Resource estimation and calculation
  - Resource pre-allocation
  - Resource brokering
  - Resource assignment

BigData Express Status

Project development & implementation has proceeded on schedule. The project team is now conducting function & performance tests:

- A BDE Web Portal prototype is available now.
- BDE DTN scheduler is planned to be released on May 2017
- BDE end-to-end network path module is planned to be released on June 2017